STUDIA PHONOLOGICA XIII (1979)

# Articulatory Effects on Speech Perception: Adaptation Test of Voicing Feature Detectors*

Katsumasa SHIMIZU

## I. Introduction

One of the basic problems in the study of speech perception is how two mechanisms of production and perception of speech are linked and related to each other. The problem may belong to the old and out-of-date question, and much discussion has been made about the relationships between two mechanisms, especially by the advocates of the motor theory. A number of studies suggest that two mechanisms are somehow related to each other and few suggest that they are completely separate mechanisms. Still, however, there has been no convincing proposal which has attained wide acceptance from various angles of investigation as to how and in what stage of speech processing they are related to each other.

Some recent approaches to this problem are directed to the function of feature detectors which are neurological cortical cells and are selectively responsive to specific features of stimuli. The studies on feature detectors suggest that repetitive exposure to a single stimulus causes the shift of phonetic boundary in the identification of speech continuum and this experimental method is called selective adaptation test (Eimas & Corbit, 1973; Ades, 1974). Cooper (1974) and Cooper & Lauritsen (1974) reported that perceptuomotor effects appeared after perceptual adaptation to voiceless stop; that is, immediately after listeners were presented with many repetitions of voiceless stop consonant, the value of voice onset time (VOT) decreased consistently as they uttered voiceless CV sequences. Furthermore, Cooper, et al. (1975) and Cooper, et al. (1976) reported that repetitive articulation of a sequence of the syllables caused a shift of phonetic boundary in the identification of some but not all speech continuums[1]. The direction of boundary shift can be considered as

---

Katsumasa SHIMIZU (清水克正): Instructor, Dept. of Linguistics, Kyoto University, and also a faculty member of Nagoya Gakuin Univ. The author is doing research in linguistics under the direction of Dr. Tatsuo Nishida, Professor of Linguistics, Kyoto Univ.

* This paper was presented at the 9th International Congress of Phonetic Sciences, Copenhagen, Denmark, August 6–11, 1979.

1. Cooper, et al. (1975) asked subjects to utter a sequence of [bæ, mæ, væ] before the identification test of synthetic speech continuum [bæ–dæ–gæ]. The identification test was carried out after one-min. periods of repetitive whispered articulation with no auditory feedback, repetitive "thinking" of the sequence, and repetitive articulation with auditory feedback. The results were less convincing to determine articulatory effects on speech perception, though some subjects showed significant effects. Cooper, et al. (1976) asked subjects to utter [si–su], [sti–stu] before the identification test of synthetic speech continuum of [si–sti], and to utter [wa–ya], [ba–da] before the identification test of synthetic speech continuum of [ba–wa]. Articulatory effects appeared in the cases of [sti–stu] and [wa–ya].

follows: when subjects utter voiced stop consonants (as CV syllable), they give more identification responses belonging to the voiceless category; conversely when they utter voiceless stop consonants, they give more identification responses belonging to the voiced category. It is explained that perceptuomotor effects and articulatory effects on speech perception are attributable to the fatigue of a mechanism that mediates production and perception of speech. However, experimental evidences are still scarce to support the direct and indirect relationships between two mechanisms in terms of feature detectors. From a rather different point of view, Bell-Berti, et al. (1978) examined whether the differences in vowel production of /i–I/ and /e–E/ correspond to the ones in perception, and they suggest a possible link between strategies in production and perceptual analysis of speech sounds[2].

In our previous study (Shimizu, 1977), articulatory effects on speech perception were examined on voicing features for labial, alveolar, and velar sets. Subjects were instructed to utter repetitively [ba] and [pa] for labial set, [da] and [ta] for alveolar set, and [ga] and [ka] for velar set. The shifts of phonetic boundaries in identification were observed in the articulation of such segments as [ba], [pa], and [ka], and it was pointed out that there were some differences in the strength of articulatory effects, depending on the differences of place features and the distinction of voiced-voiceless features; articulatory effects were more consistent in voiceless feature than in voiced one, and labial features were more sensitive than other place features. It has become clear through the experiment that repetitive articulation of some CV syllables produced a shift of phonetic boundaries; that is, some articulatory effects were observed on speech perception. We felt it necessary, however, to further examine whether the detectors operate on units of phonetic features or on further smaller units, and to carry out the cross-series adaptation experiments. Under such circumstances, the present study aims at further developing the understanding on feature detectors and at examining how repetitive articulation in the cross-series experiment affects the mechanisms of speech processing.

## II. Experimental Method

1) Subjects: Ten Japanese speakers took part in the experiment as a subject, ages from 18 to 22. All subjects were undergraduate students of Nagoya Gakuin University. They had normal hearing ability and had no known neurological defects.

2) Stimuli: The stimuli for identification consist of two sets; alveolar and velar sets. They are the same with the ones used in the experiment carried out in 1977. They are three-formant synthetic speech sounds prepared by a parallel resonance synthesizer at Haskins Laboratories. The alveolar set consists of 24 variants of

2. Bell-Berti, et al. revealed two different ways to differentiate the members of pairs /i–I/ and /e–E/. The one is based on tongue height, while the other is the degree of tongue tension. They examined whether these differences in vowel production may correspond to the ones in perception.

/da–$t^{h}$a/ series with the VOT continuum from −40 to +80 msec. The velar set consists of 25 variants of /ga–$k^{h}$a/ series with the VOT continuum from −50 to +70 msec. The variations in VOT were made by varying the onset of the first formant relative to the onset of higher formants. The stimuli in each set vary in 5 msec steps.

3) Procedure: The experiment included unadapted identification test of alveolar and velar sets, identification test of the same sets after one-minute periods of repetitive articulation, and measurement of voice onset time.

To obtain the unadapted identification functions, 24 stimuli in alveolar set and 25 stimuli in velar set were presented binaurally at a comfortable listening level. For each set, 6 times of repetition of each stimulus were presented in a randomized order and the interval between stimuli was about 5 seconds. Next, to obtain the cross-articulatory effects on speech perception, each subject was instructed to utter [ga] and [ka] for alveolar set and [da] and [ta] for velar set. Subjects were tested one at a time in a sound-proofed booth. They were instructed to utter [ka] for alveolar set and after half an hour later to utter [ta] for velar set. Each subject uttered voiceless stop consonants for one minute repetitively and immediately after this, ten test stimuli were presented for identification, and this process was repeated until all test stimuli in each set were presented. On a separate day, they were instructed to utter voiced stop consonants as in the same way as voiceless stops[3]. They uttered voiced stop consonants repetitively for one minute and immediately after this, they did the identification test of ten test stimuli, and this process was repeated until all test stimuli in each set were presented. Repetitive articulations of all subjects were recorded by Sony tape recorder TC-707MC.

To measure voice onset time, the initial (first) and final (last) utterances of each subject were recorded and analyzed by Kay Sonagraph. The wide-band spectrograms for such utterances were prepared.

## III. Results

As we did in our previous study (Shimizu, 1977), the phonetic boundaries of the voiced and voiceless stop consonants for alveolar and velar sets were extrapolated for each subject from the responses in the identification test[4]. The boundary was defined as that point on the stimulus scale which would receive 50% responses from voiced and voiceless categories in each set. The results are shown in Table 1.

What we are interested in this experiment is whether the locus of phonetic boundary shifts when identification stimuli and the articulating syllables are from different series. The boundary shift is the difference between base boundary and the boundary after repetitive articulation. Plus value indicates that the boundary is shifted to right direction, while minus value indicates that the boundary is shifted

3. We carried out the experiments for voiced and voiceless stop consonants on a separate day for each subject so that any effects from articulation of either of consonants could be eliminated.
4. The method to extrapolate the phonetic boundary was adopted from Donald (1976).

Table 1. Shift of the phonetic boundary in milliseconds of VOT in the cross-series experiment of repetitive articulation.

| Subjects | Alveolar Set | | | Velar Set | | |
|---|---|---|---|---|---|---|
| | Base [ta-da] boundary | Shift after [ka] articu. | Shift after [ga] articu. | Base [ka-ga] boundary | Shift after [ta] articu. | Shift after [da] articu. |
| AS | +18 msec | −3 | +12 | +18 msec | +7 | −10 |
| SK | +26 | 0 | −3 | +23 | −1 | −3 |
| HS | +20 | +10 | 0 | +25 | +3 | −3 |
| MO | +19 | +5 | +11 | +25 | +5 | 0 |
| CY | +32 | −2 | +1 | +32 | +1 | −5 |
| TF | +25 | −1 | 0 | +23 | +4 | +4 |
| HI | +24 | +6 | +2 | +22 | +3 | —— |
| UT | +20 | +15 | +10 | +26 | +4 | +1 |
| ST | +25 | +2 | −2 | +30 | +3 | −4 |
| MF | +20 | +4 | +4 | +13 | +5 | +5 |

to left direction. From Table 1, in alveolar set, subjects HS, MO, HI, UT, and MF showed a relatively significant shift in the articulation of voiceless stop [ka], as predicted in the perceptual adaptation of voiceless stops, but the articulation of voiced stop [ga] did not show any predictable shift. Rather, it caused the same direction of shift as seen in the articulation of voiceless stop [ka]. We should note that subjects MO, HI, UT, and MF showed the same direction of boundary shift, regardless of voiced-voiceless distinction of the articulating syllables. In velar set, most subjects except SK showed a shift of phonetic boundary to predicted direction in the articulation of voiceless stop [ta], though the effects of subjects HS, CY, HI, and ST were not so significant. In the overall examination, it can be pointed out

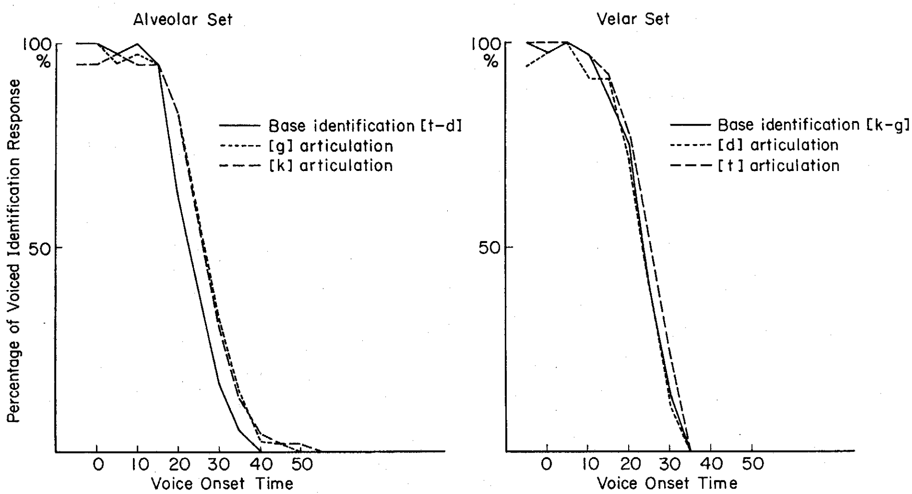


Fig. 1. Percentage of voiced identification responses [d] and [g] obtained with and without articulatory effects. The solid lines indicate the base identification functions and the dotted and dashed lines, the identification functions after repetitive articulation of voiced and voiceless consonants.

Table 2. VOT values in repetitive articulation of vioceless stop consonants.

| Subjects | [ta] articulation Initial utterance | [ta] articulation Final utterance | [ka] articulation Initial utterance | [ka] articulation Final utterance |
|---|---|---|---|---|
| HS | +21.0 msec | + 8.4 msec | +42.0 msec | +54.6 msec |
| AS | +12.6 | +16.8 | +29.4 | +16.8 |
| ST | +21.0 | +16.8 | +58.0 | +46.2 |
| MF | +16.8 | +16.8 | +29.4 | +41.2 |
| SK | +21.0 | + 8.4 | +50.4 | +47.9 |
| MO | +16.8 | +21.8 | +42.0 | +26.9 |
| HI | +42.0 | +25.2 | +25.2 | +31.1 |
| Mean | +21.6 | +16.3 | +39.5 | +37.8 |

that the cross-series articulatory effects were significant in the articulation of voiceless stops, but not in voiced stops. This may imply that the detectors which are considered to mediate two mechanisms function at "distinctive feature" level, not as phonetic unit level. In Figure 1, the cross-articulatory effects are shown in curves for alveolar and velar sets.

Table 2 indicates the voice onset time values of voiceless stop consonants in repetitive articulation. We reasoned to measure the VOT values as follows: if the direct effects from repetitive articulation come out, it would be the duration of voice onset time that would be affected. If this is the case, repetitive articulation of CV syllables would affect the neural commands which control the release of articulators. The measurement of the VOT values was made under the criterion described by Lisker and Abramson (1964); that is, "the point of voicing onset was determined by locating the first of the regularly spaced vertical striations, while the instant of release was found by fixing the point where the pattern shows an abrupt change in overall spectrum" (1964: 389). Wide-band spectrograms were prepared from the recordings of initial and final parts of utterances of [ta] and [ka]. Initial utterances were taken from the first one-minute utterances, while final utterances were taken from the tenth one-minute utterances. In most cases, subjects uttered CV syllables about 65 times in a minute.

The VOT values of voiceless stop consonants [ta] and [ka] seem to belong to the category of voiceless unaspirated stop according to three categories described by Lisker and Abramson (1964), and, as predicted, the mean value of velar stop was longer than that of alveolar stop. In examining the VOT values in both utterances, the data is rather less-than-convincing concerning systematic effects produced by repetitive articulation; however, as an overall tendency, repetitive articulation produced a decrease of the VOT values. Four out of seven subjects showed the decrease of the values, as they uttered repetitively.

## IV. Discussion

In examining the results of the present experiment, we should note that the shift of phonetic boundary occurred even in the case when the identification stimuli and the repetitively articulated syllables were from different series. Repetitive articulation of voiceless stop consonants in the cross-series experiment caused a shift of the boundaries in alveolar and velar sets, the boundaries being displaced toward the voiceless category. On the other hand, the repetitive articulation of voiced stop consonants did not show any consistent and predictable direction of shift. The cross-series effects in voiceless stop consonants imply that the detectors for voicelessness function as a phonetic feature, not as a unit. The reasoning for the effects is that the repetitive articulation of voiceless stop consonants should fatigue the function of the detector and should reduce the sensitivity of the detector in perception. The fatiguing of the detector in articulation leads to the variation of the VOT values. In our present study, the VOT values tended to be shortened as subjects uttered voiceless stop consonants repetitively, though the data was less convincing. We can infer that the shortened VOT values in articulation may lead to more responses belonging to voiced category in identification, resulting the shift of phonetic boundary toward voiceless category. That is, some stimuli which were identified as voiceless in unadapted test were perceived as voiced after the articulation.

The cross-series effects have some important implications in the study of speech processing. First, the effects imply the existence of phonetic feature detectors and may give concrete support for the description of distinctive feature system[5]. Second, the detectors may function as a common mechanism which mediates perception and production of speech. Third, the levels in which the detectors function, especially for voiceless feature, are not peripheral in speech processing.

We have no plausible explanation as to why different effects appeared between voiced and voiceless stop consonants in the cross-series experiments. Cooper & Lauritsen (1974) reported the similar results when they examined the perceptuo-motor effects; the effects appeared in voiceless Plosives, but not in voiced plosives. One of the differences between voiced and voiceless stop consonants is the distributional pattern of the VOT values. They reported that the values for voiceless stop consonants vary in a wide range from +20 to +110 msec, while the ones for voiced stop consonants vary in a narrow range from 0 to +20 msec. This may indicate that voiced and voiceless features may receive different neural commands in articulation, possibly in controlling the release of articulators, and that they are separately processed in speech perception.

## V. SUMMARY

1) In the articulation of voiceless stop consonants, the shift of phonetic boundary occurred even in the case when the identification stimuli and the articulating

5. Distinctive feature system implies here the one proposed by Jakobson, R., Fant, G. and Halle, M. (1963).

syllables were from different series, but the repetitive articulation of voiced stop consonants did not show any systematic effects. The cross-series effects of voiceless stop consonants imply that the detectors are selectively sensitive to voiceless feature and they function as a phonetic feature, not as a unit.

2) Repetitive articulation of voiceless stop consonants caused the VOT values to shorten in some subjects. The shortened VOT values in articulation may cause the shift of phonetic boundary, the boundary being displaced toward voiceless category.

## Acknowledgements

I wish to thank Dr. Tatsuo Nishida of Kyoto Univ. who has assisted and encouraged my study in linguistics in various ways. I also wish to thank Dr. L. Lisker, Dr. A. S. Abramson and Dr. C-W. Kim who gave comments on this paper, when I presented it at the Ninth International Congress of Phonetic Sciences, August 6–11, 1979, in Copenhagen, Denmark.

## References

Abramson, A. S. (1976), "Laryngeal timing in consonant distinctions", Haskins Laboratories, *Status Report on Speech Research*, 47, 105–112.

Ades, A. E. (1974), "How phonetic is selective adaptation? Experiments on syllable position and vowel environment", *Perception and Psychophysics*, Vol. 16(1), 61–66.

Bell-Berti, F., Raphael, L. J., Pisoni, D. B. and Sawusch, J. R. (1978), "Some relationships between articulation and perception", Haskins Laboratories, *Status Report on Speech Research*, 55/56, 21–32.

Cooper, W. E. (1974), "Perceptuomotor adaptation to a speech feature", *Perception & Psychophysics*, Vol. 16(2), 229–234.

Cooper, W. E. and Lauritsen, M. R. (1974), "Feature processing in the perception and production of speech", *Nature*, Vol. 252, 121–123.

Cooper, W. E., Blumstein, S. E. and Nigro, G. (1975), "Articulatory effects on speech perception: a preliminary report", *Journal of Phonetics*, 3(2), 87–98.

Cooper, W. E., Billings, D. and Cole, R. A. (1976), "Articulatory effects on speech perception: a second report", *Journal of Phonetics*, 4(3), 219–232.

Darwin, C. J. (1976), "The perception of speech", *Handbook of Perception*, Vol. VII, (ed. by E. C. Carterette and M. P. Friedman), Academic Press, 175–226.

Donald, S. L. (1976), "Selective adaptation of labial stops for Thai and English speakers", a paper presented at LSA Annual Meeting, December 30, 1976.

Eimas, P. D. and Corbit, J. D. (1973), "Selective adaptation of linguistic feature detectors", *Cognitive Psychology*, 4, 99–109.

Lane, H. (1965), "The motor theory of speech perception: a critical review", *Psychological Review*, Vol. 72(4), 275–309.

Lisker, L. and Abramson, A. S. (1964), "A cross-language study of voicing in initial stops: acoustical measurements", *Word*, 20, 384–422.

Shimizu, K. (1977), "Voicing features in the perception and production of stop consonants by Japanese speakers", *Studia Phonologica*, XI, 25–34.

Studdert-Kennedy, M. (1976). "Speech perception", *Contemporary Issues in Experimental Phonetics* (ed. by N. J. Lass), Academic Press, 243–293.

Studdert-Kennedy, M. (1979), "Speech perception", *Proceedings of Ninth International Congress of Phonetic Sciences*, Vol. 1, Copenhagen, 59–81.

(Aug. 31, 1979, received)